



FACTORS CONTRIBUTING TO TIME SPENT ON PAID AND UNPAID WORK: A REGRESSION ANALYSIS OF TIME USE SURVEY DATA

Suzie Carson

*Social Policy
Statistics New Zealand*

Frances Krsinich

*Analytical Support
Statistics New Zealand*

Susan Kell

Ministry of Women's Affairs

Abstract

The study of time spent on different activities rather than the study of people can offer new insights as well as new challenges. This paper presents the results from an exploratory analysis of data from the first New Zealand Time Use Survey. Interesting questions arise about the distribution of paid and unpaid work across different groups in society. Of particular interest is how men and women balance the competing demands of paid and unpaid work. This paper is a first attempt at understanding some of the complex relationships between social, demographic and employment characteristics and paid and unpaid work time.

Keywords: *Paid work, unpaid work, Time Use Survey*

The New Zealand Time Use Survey was conducted by Statistics New Zealand under contract to the Ministry of Women's Affairs. The data collection was conducted over a twelve-month period from July 1998 to June 1999. The survey population is defined as the civilian, usually resident, non-institutionalised population aged 12 years and over residing in private households. The survey instruments comprised a 48-hour diary, a personal questionnaire and a household questionnaire. The sample was allocated evenly across the 12 months of the survey period to minimise the seasonal effects and was also balanced across days in the week. The survey had a response rate of 72 percent corresponding to an achieved sample size of 8,522 respondents. The sample selection procedure included a separate Maori screening sample to boost the Maori sample size and improve the reliability of Maori population estimates. Up to two people were randomly selected within each household. Each respondent in the survey was assigned a unique survey weight that adjusted for the probability of selection, non-response and calibrated the sample totals to the population benchmarks.

Further details of the survey methodology are available from Statistics New Zealand.

Regression Analysis Methodology

Multiple regression modelling was used to explore the factors that explain the time spent in paid work and total unpaid work. The paid and unpaid work models were estimated using a general linear model using PROC GLM in SAS. With the exception of work hours the explanatory variables were treated as categorical variables. The paid and unpaid work models were initially estimated using the corner point method however the large number of explanatory variables makes the interpretation of the corner point estimates difficult. The population marginal means were also estimated because these give an average value for each population group in the model and are therefore easier to interpret.

The original model specifications contained a large number of possible explanatory variables. The full model

was estimated and variables with a P-value of 0.05 or higher in the type III sums of squares were removed from the model. Initially, the dataset was randomly split into two equally sized datasets and the initial model selection was undertaken on the first half dataset. These fitted models were then run on the second half dataset to check the robustness of the model before the final models were estimated using the full dataset.

A number of methodological issues arise in the use of time use data for regression analysis of this type. The New Zealand Time Use Survey collected diary information in five minutes intervals from respondents for two concurrent 24-hour days. These two diary days cannot be treated as independent observations and limiting the analysis to one diary day would result in a loss of information. Instead, the amount of time spent on the relevant activity was summed over the two diary days and this was used as the dependent variable. This limits the analysis to respondents who completed two diary days (99.7 percent of the total sample).

In the time use survey both primary and simultaneous activities were collected in the diary. However, this analysis is restricted to primary activities where primary activity is defined (in most cases) as the activity recorded in the first column of the diary.

The Time Use Survey diaries were collected for both weekdays and weekend days. It is anticipated time use would vary between weekdays and weekend days. For this reason the combination of the days of the week surveyed was included as an explanatory variable in the regression models.

One limitation of using a cross-sectional dataset like the Time Use Survey for regression analysis is that it is not possible to determine the direction of causation. In this analysis we have considered the demographic, social and economic variables that are associated with different time use patterns. In some cases however, it may be that the time use influences an explanatory variable, for example, low income may be a consequence of few hours spent in paid work and not vice versa.

The paid and total unpaid work models were tested for heteroscedasticity. Heteroscedasticity occurs when the assumption that the variance of the error term is constant across all observations is violated. The consequence of heteroscedasticity is that the estimators are unbiased but the standard errors are incorrectly estimated. There was evidence of heteroscedasticity in both the paid and unpaid work models. As a consequence the estimated standard errors may under or over estimate the true error. The estimation procedure did not allow for calculation of heteroscedastic corrected standard errors. In addition, the statistical package used in this analysis did not account for the complex sample design and so although a nominal 95 percent significance level was employed it is likely that the true coverage probability is less than 95 percent. The base category for both the paid and total unpaid work

models was non-Maori females aged 45-54 years with no children (aged less than 18 years), part-time employed (people who are not in the labour force for unpaid work model), did not pay for housework or childcare, had two weekday diary days, income of between \$10,000 and \$40,000 and lived in an urban area. The allocation of the base categories for the explanatory variables is an arbitrary decision and does not affect the significance of the variables. The base categories were chosen because they were the most meaningful categories on which to base comparisons for each explanatory variable.

Both corner point and population marginal means were estimated. The corner point estimates are interpreted in relation to the base category. Data that is cross-classified (i.e. having observations of the response variable in cells defined by combinations of the classes corresponding to the explanatory variables) can be modelled using a linear model with categorical regressors. Typically such linear models are over specified in the sense that there is a linear relationship between groups of the categorical regressors. There are several ways to remove these redundancies. A useful way for unbalanced data is the "corner point" or "set to zero" constraint. Consider a model with one categorical regressor A with classes 1, 2, and 3. The data could be modelled directly via parameters corresponding to the cell means μ_1 , μ_2 , and μ_3 . The corner point parameterisation would instead take one cell mean say μ_1 as a parameter, and then set the other parameters to be the difference between this cell mean and the remaining cell means, so the set of parameters are: μ_1 , $\mu_2 - \mu_1$, $\mu_3 - \mu_1$. An analysis of variance on this linear model is used to determine which effects or which interactions between effects are significant.

The corner point estimates are of limited use in a model of this kind with many explanatory variables and a complex base category. A researcher is may also be interested in estimating the means (expected values) of particular effects. The means reported in this paper are population marginal means and were calculated using an LSMEANS statement in SAS. The population marginal mean is a linear combination of the model parameters averaged over specified classes as if there was only one observation per cell. This definition doesn't depend on the sample sizes in the cells, which could vary in unbalanced designs.

The population marginal means are estimated using the fitted model and make an adjustment for the unbalanced survey data. Balanced or experimental data is collected in such a way to allow for all categories of the explanatory variables to have the same number of observations. The time use survey data was not collected using a balanced design, but the estimates of the population marginal means make an adjustment for this unbalanced design. The population marginal means for any specific group are only helpful in the context of these models and should be interpreted with caution in light of the low R-squared statistics that were obtained for both models. This is reflected in the comparatively high standard errors that

were obtained for the population marginal means. While accurate cell means are available from the time use survey these are less helpful in understanding the complex interrelationships between variables. The population marginal means reported in this paper will differ from the averages reported elsewhere from the time use survey.

What Factors Influence Paid Work Time?

In this analysis demographic, family and employment characteristics are examined to determine their relevance in explaining paid and unpaid work time differences. Paid work time includes work for pay or profit, education or training in work time, job search activities, travel associated with labour force activity and other labour force activity not elsewhere classified. The paid work model identifies the factors that explain paid work time for people who are employed for at least one hour per week. Therefore, the model identifies the factors that explain variations in time spent in paid work rather than the factors related to whether a person is employed or not. The paid work model includes both main effects and relevant interactions between the explanatory variables and includes nested interactions where one variable is limited to a subset of categories of the other variable, for example, age of youngest child is only relevant to people with children. The number of children and age of youngest child variables only relate to people with a youngest child aged less than 18 years. The full model was estimated and those variables with a P-value of 0.05 or greater in the type III sums of squares were removed. Table 1 lists the variables that were removed because they were not found to explain time spent in paid work for people who spent one or more hours per week in paid work.

Occupation, industry and the number of jobs were not significant factors in influencing the amount of time spent in paid work. The number of children and the age of the youngest child were also not relevant in explaining differences in the amount of time spent in paid work. No differential effect was observed for men and women by the number of children. Educational qualifications and Maori ethnicity were also not found to explain differences among people in paid employment. The influence of qualifications, employment status and paying for childcare on paid work time did not differ for men and women. Partnership status had a P-value of <0.05 in the type III sums of squares in the original model but it was not significant in the final model.

The parameter estimates in table two are reported in hours per week for the cornerpoint estimates and the population marginal means. The corner point estimate of 23 hours per week for the full time category is interpreted to mean that people with the same characteristics as the base group, but who worked full-time, spent 23 hours per week more in paid work than the base group.

The paid work model for employed people yields results that are not surprising; labour force status, age group and day of the week are significant predictors of the amount

Table 1. Non-significant Variables Removed from the Paid Work Model

Total paid work
• Paid for housework
• Highest qualification
• Ethnic group
• Number of jobs
• Number of children
• Occupation
• Industry
• Sex
• Urban/rural residence
• Partnership status
• Sex and employment status
• Sex and age of youngest child
• Sex and paid for childcare
• Sex and highest qualification
• Sex and number of children
• Age of youngest child nested in number of children
• Paid for childcare nested in number of children

of time spent in paid work for employed people. Although we know the variables themselves are significant some of the reported differences from the population marginal means may not be significant but it is interesting to look at the emerging patterns. For people working one or more hours per week, people aged 12-17 years and 65 years or over spent the fewest hours in paid work. The average amount of time spent in paid work was found to be fairly stable for people aged 18-64 years.

Employment status and income were also significant factors in explaining paid work time. The paid work model revealed that on average self-employed people worked longer hours per week than paid employees and those working unpaid in a family business. The amount of time spent in paid work also differed for people in different income groups. Employed people with income of \$10,000 or less spent less time in paid work than people with an income of \$10,001-\$40,000 and an income of \$40,001 or more.

Somewhat unexpectedly neither the number of children nor the age of the youngest child were found to be significant factors in explaining the paid work time of either men or women. This finding could be due to uncontrolled heteroscedasticity, the complex sample design or because the model was unable to account for what are likely to be complex relationships between the variables.

What Factors Influence Total Unpaid Work Time?

The second model considers the factors that are associated with total unpaid work time. An activity is classed

Table 2. Corner Point Estimates and Population Marginal Means from the Paid Work Model - In Hours per Week

Variables	Corner Point parameter estimates	Standard errors	Population marginal means	Standard Errors
Intercept	** 29	1.1	-	-
Age group				
12-17 years	** -7	1.8	24	3.1
18-24 years	* 3	1.3	34	2.9
25-34 years	0	0.9	31	2.8
35-44 years	-1	0.9	30	2.8
45-54 years	31	2.8
55-64 years	0	1.2	30	2.9
65 years or over	* -7	2.2	24	3.5
Days of the week				
Two weekend days	** -30	0.9	15	2.9
Weekend and weekday	** -17	0.7	28	2.8
Two weekdays	45	2.8
Labour force status				
Full-time/part-time not specified	** 8	1.2	27	2.9
Full-time employed	** 23	0.9	42	2.8
Part-time employed	19	2.8
Employment status				
Self-employed	** 3	0.8	28	1.2
Working unpaid in a family business	** -9	2.1	16	2.3
Paid employee	25	1.1
Income				
Less than \$10,001	** -5	1.1	25	2.8
\$10,001-\$40,000	30	2.8
Over \$40,000	0	0.8	31	2.9
Sex and paid for housework				
Male paid for housework	** 7	1.7	31	3.1
Male did not pay for housework	** 6	0.7	30	2.7
Female paid for housework	0	1.9	24	3.1
Female did not pay for housework	24	2.7

$R^2 = 0.39$

'...' denotes the non-estimable base-categories under the corner point estimation

* P-value ≤ 0.05

** P-value ≤ 0.0001

as unpaid work if it is a productive activity that has no remuneration and satisfies the third person criterion, that is, the activity yields an output that can be exchanged. Total unpaid work time includes unpaid work for people living in the same household and for people in other households. Unpaid work includes activities such as household work, caregiving for household members, purchasing goods and services for your own household and unpaid work for people outside of your own household. The model includes people who did not do any unpaid work in the two diary days surveyed as well as those doing unpaid work. Therefore the model includes both the factors that influence participation in unpaid work as well as differences in the time spent in unpaid work activities. Table 3 shows the non-significant variables that were removed from the total unpaid work model.

As Table 4 shows, sex was a significant factor in explaining differences in unpaid work time; males aged 12 years or over spent far fewer hours per week on unpaid work when compared to females. The age of the youngest child was not significant when nested with number of children but the effect of the age of the youngest child on unpaid work was found to differ between men and women. Sex differences in unpaid work time were also apparent by age group and labour force status. Partnered and non-partnered males spent similar amount in unpaid work, as did partnered and non-partnered women but there was a significant sex difference with partnered and non-partnered women spending more time doing unpaid work.

Age group was a significant factor in explaining unpaid work time with a generally increasing trend towards more unpaid work as age increased. The amount of time spent

Table 3. Non-significant Variables Removed from the Total Unpaid Work Model

Total unpaid work
• Whether received a benefit
• Ethnic group
• Number of children
• Partnership status
• Income level
• Urban/rural
• Sex and paid for housework
• Sex and paid for childcare
• Sex and income level
• Sex and number of children
• Paid for childcare for people who had children
• Age of youngest child for people who had children
• Number of jobs for employed people
• Occupation for employed people
• Industry for employed people
• Employment status for employed people

in unpaid work for people of the same age group also differed by sex. In general females spent more time in unpaid work than males in the same age group.

Labour force status was a factor in unpaid work hours, overall those in full-time and part-time employment spent the fewest hours in unpaid work and unemployed people and those not in the labour force the most. However, the effect of labour force status was modified by sex with women spending more time in unpaid work on average for the same labour force status. Women who were unemployed or not in the labour force spent the most hours in unpaid work but women in full-time or part-time employment still retained significant unpaid work responsibilities.

The number of children did not emerge as a significant factor in explaining unpaid work time. However, the age of the youngest child was a significant factor in determining the number of hours spent on unpaid work and this effect was different for men and women. While men with a youngest child aged less than 5 years spent more time in unpaid work activities than men with no children, women with a youngest child aged less than 5 years spent the most time in unpaid work on average. Unpaid work time was highest for people with children aged less than 5 years, while unpaid work time was similar for people with no children and people with a youngest child aged between 14 and 17 years.

Discussion and Policy Implications

Time use data can be used to inform a range of policy areas such as employment, health, education, welfare, transport, civil defence as well as gender and ethnicity differences among population groups and estimates of

national production. The time use data can also better inform policy making by providing contextual information about people's lives. The multiple regression analysis was confined to the factors associated with paid and unpaid work time for people with differing characteristics in order to explore the inter-relatedness of paid and unpaid work activities.

Age group was found to be a significant variable in explaining both paid and unpaid work time, as was the day of the week and labour force status. Sex in itself was not found to be significant in explaining paid work time of employed people but was very significant in the unpaid work model. Sex also modified the effect of a number of other factors in the unpaid work model. Employment characteristics like occupation, industry and number of jobs were not significant in explaining paid or unpaid work time. Family characteristics like the age of the youngest child and partnership status were found to explain unpaid work time and in some cases the effect of these family characteristics on unpaid work time were modified by sex.

One finding of particular interest is the large difference in the unpaid work time of males and females in the same age group. This large difference would suggest consequences for women's time available for employment, education, physical exercise and hobbies. The larger share of unpaid work among women may also affect educational outcomes over their lifetime, and their opportunities to acquire interests and skills that would be relevant to future employment. One policy response may be to educate young people to share the unpaid work responsibilities more equally with their partners in adult life.

Another emerging theme was the large differences in unpaid work responsibilities between men and women with similar circumstances, for example, those in the same labour force status and for those in families where the youngest child is in the same age group. While women in full-time and part-time employment do fewer hours unpaid work than unemployed women or women who are not in the labour force, they retain significant unpaid work responsibilities. The reality of the uneven distribution of unpaid work between women and men is an issue for employment and childcare policies. Employers could make greater efforts in providing family friendly workplaces for people with family responsibilities. In addition to the provision of childcare, employers could improve the provision of workplace parking, provide access to household laundry and dry-cleaning services, provide care for children who are out of school because of illness or after school supervision of older children.

Future Research

Future research in the area of paid and unpaid work might involve quantifying the relative importance of the family, employment and demographic characteristics in explaining differences in time use and investigating the trade-offs between different types of time use, for example, how

Table 4. Corner Point Estimates and Population Marginal Means from the Unpaid Work Model - In Hours per Week

Variables	Corner Point parameter estimates	Standard errors	Population marginal means	Standard errors
Intercept	** 42	1.4	-	-
Age group				
12-17 years	** -23	1.1	16	3.7
18-24 years	** -14	1.1	22	3.7
25-34 years	** -7	1.0	29	3.7
35-44 years	* -4	0.9	33	3.7
45-54 years	36	3.7
55-64 years	* 4	1.1	39	3.7
65 years or over	* -3	1.1	36	3.7
Days of the week				
Two weekend days	** 4	0.5	32	3.7
Weekend and weekday	** 2	0.4	30	3.7
Two weekdays	28	3.6
Paid for housework				
Yes	* -2	0.7	30	3.7
No	32	3.6
Highest qualification				
No qualification	* -1	0.5	28	3.7
School qualification	31	3.7
Post-school only	* 2	0.7	31	3.7
Post-school and school	* 1	0.5	30	3.7
Labour force status				
Full-time employed	** -17	0.7	21	3.6
Part-time employed	** -5	0.8	27	3.8
Unemployed	0	1.4	36	4.3
Not in the labour force	37	3.7
Sex				
Male	** -16	2.1	21	5.8
Female	40	4.2
Sex and partnership status				
Male non-partnered	0	0.7	24	1.2
Male partnered	24	1.1
Female non-partnered	** -4	0.6	34	1.1
Female partnered	38	1.1
Sex and labour force status				
Male full-time employed	** 4	1.1	12	5.8
Male part-time employed	2	1.3	18	6.0
Male unemployed	3	2.0	27	6.3
Male not in the labour force	27	5.9
Female full-time employed	29	4.2
Female part-time employed	37	4.3
Female unemployed	45	4.7
Female not in the labour force	47	4.3
Sex by age of youngest child				
Male and no children	1	1.5	16	5.8
Male and youngest less than 5 yrs	** 14	1.6	29	5.9
Male and youngest aged 5-13 yrs	** 8	1.5	23	5.9
Male and youngest aged 14-17 yrs	15	6.0
Female and no children	-2	1.3	29	4.2

women with children balance the competing demands of paid and unpaid work. This analysis deals with only one aspect of the potential policy uses of the time use survey. The time use survey offers potential for investigating a wide range of issues. The time use survey could be used to explore the following research and policy issues: recognising the role of unpaid work in the care of older adults, understanding the range and extent of unpaid productive activities undertaken by people in receipt of government benefits, identifying the whereabouts of the population at particular times of the day to assist civil defence and transport policies and exploring participation in cultural and community life.

Notes

1. The authors would like to thank Alistair Gray for his assistance and advice on the multiple regression modelling.
2. The findings presented in this paper are those of the authors and are not necessarily the official view of Statistics New Zealand.

References

- Searle, S. R. Speed, F. M. and Milliken, G. A. (1980)
Population Marginal Means in the Linear Model:
An Alternative to Least Squares Means. *The
American Statistician*, 34 (4): 216-221

Authors

Suzie Carson
Social Statistics Analyst
Social Policy
Statistics New Zealand
PO Box 2992
Wellington
Suzie_Carson@stats.govt.nz

Frances Krsinich
Mathematical Statistician
Analytical Support
Statistics New Zealand
PO Box 2992
Wellington
frances_Krsinich@stats.govt.nz

Susan Kell
Senior Policy Analyst
Ministry of Women's Affairs
PO Box 10 049
Wellington
kell@mwa.govt.nz